VERIFICATION OF TRANSLATION

I, Kazunori Hashimoto

of 6th Fl., Kantetsu Tsukuba-Science-City Bldg. 1-1-1, Oroshi-machi, Tsuchiura, Ibaraki, JAPAN

declare as follows:

1. That I am well acquainted with both the English and Japanese languages, and

2. That the attached document is a true and correct translation of the certified copy made by me to the best of my knowledge and belief of:-

(a) Japanese Patent Application No. Hei 8-185216
Entitled: " NOVEL VEGF-LIKE FACTOR "
Filed on July 15, 1996

November 22, 2001 Kazunori Hashimoto

(Date)

(Signature of Translator)
Kazunori Hashimoto
Patent Attorney

[Name of Document] APPLICATION FOR PATENT

[Identification Number] C1-802

[Filing Date] July 15, 1996

[Addressee] Esq. Commissioner of the Patent Office

[International Patent Classification] C12N 15/18

[Title of the Invention] NOVEL VEGF-LIKE FACTOR

[Number of Claims] 5

[Inventor]

[Address] c/o CHUGAI RESEARCH INSTITUTE FOR MOLECULAR MEDICINE, INC.
153-2, Nagai, Niihari-mura, Niihari-gun, IBARAKI

[Name] Yuichi HIRATA

[Inventor]

[Address] c/o CHUGAI RESEARCH INSTITUTE FOR MOLECULAR MEDICINE, INC.
153-2, Nagai, Niihari-mura, Niihari-gun, IBARAKI

[Name] Junichi NEZU

[Applicant]

[Zip code] 300-41

[Address] 153-2, Nagai, Niihari-mura, Niihari-gun, IBARAKI

[Name or Appellation] CHUGAI RESEARCH INSTITUTE FOR MOLECULAR MEDICINE, INC.

[Representative] Tsutomu KAWAGUCHI

[Attorney]

[Identification Number] 100102978

[Patent Attorney]

[Name or Appellation] Hatsushi Shimizu

【Payment】

【Number of Register】 041092

【Amount】 21000

【List of Filing Papers】

【Name of Item】 Specification 1

【Name of Item】 Drawings 1

【Name of Item】 Abstract 1

【Proof】 Necessary

[Document Name] Specification

[Title of the Invention] Novel VEGF-Like Factor

[Claims]

[Claim 1] A DNA described in SEQ ID NO: 2.

[Claim 2] A DNA hybridizing with the DNA described in SEQ ID NO: 2.

[Claim 3] A vector containing the DNA of Claim 1 or 2.

[Claim 4] A transformant carrying the vector of Claim 3.

[Claim 5] A protein encoded by the DNA of Claim 1 or 2.

[Detailed Description of the Invention]

[0001]

[Field of the Invention]

The present invention belongs to the field of genetic engineering, in particular, a protein factor involved in angiogenesis in humans.

[0002]

[Prior Art]

The process of angiogenesis, in which endothelial cells existing in the inner wall of blood vessels of animals generate new blood vessels, is triggered by transduction of a specific signal. A variety of substances have so far been reported to be involved in this signal transduction. The most notable substance among them is the vascular endothelial growth factor (VEGF). VEGF is a protein factor which was isolated and purified, and can increase the proliferation of endothelial cells and the permeability of blood vessels (Senger, D. R. et al., Science 219: 983-985 (1983); Ferrara, N. and Henzel, W. J., Biochem. Biophys. Res. Commun. 161: 851-858 (1989)). It has been reported that the human VEGF gene contains eight exons and produces four subtypes consisting of 121, 165, 189, or 206 amino acid residues, depending on the difference in splicing, which causes different secretion patterns (Houck, K. A. et al., Mol. Endocrinol. 5: 1806-1814 (1991)). It has also been reported that there is a VEGF-specific receptor, flt-1, and that the binding of VEGF to flt-1 is important for the signal transduction (Vries, C. D. et al., Science 255: 989-991 (1992)).

[0003]

Placental growth factor (PlGF) and platelet-derived growth

factor (PDGF) have thus far been isolated and are factors related to VEGF. These factors are found to promoteproliferation activities of vascular endothelial cells (Maglione, D. et al., Proc. Natl. Acad. Sci. USA 88: 9267-9271 (1991); Betsholtz, C. et al., Nature 320: 695-699 (1986)). In addition, VEGF-B (Olofsson, B. et al., Proc. Natl. Acad. Sci. USA 93: 2576-2581 (1996)) and VEGF-C (Lee, J. et al., Proc. Natl. Acad. Sci. USA 93: 1988-1992 (1996); Joukov, V. et al., EMBO J. 15, 290-299 (1996)) have recently been isolated.

[0004]

These factors appear to constitute a family, and this family may contain additional unknown factors.

[0005]

It has been suggested that VEGF is involved in not only vascular formation at the developmental stage but also in the pathological neovascularization associated with diabetes, rheumatoid arthritis, retinopathy, and the growth of solid tumors. Consequently, the VEGF family factors are presently drawing extraordinary attention for developing useful, novel drugs.

[0006]

[Problems to be Solved by the Invention]

An objective of the present invention is to isolate a novel gene of the VEGF family.

[0007]

[Means to Solve the Problems]

The present inventors searched for genes having homology to VEGF-C, which is a recently cloned VEGF family gene, against Expressed Sequence Tags (EST) and Sequence Tagged Sites (STS) in the GenBank database. As a result, the inventors found an EST that was assumed to have homology to the C-terminal portion of VEGF-C. We then designed primers based on the sequence, and amplified the corresponding cDNA using the 5' RACE method and the 3' RACE method. The nucleotide sequence of the amplified cDNA was determined, and the deduced amino acid sequence therefrom revealed that the amino acid sequence had high homology to that of VEGF-C. Based on the homology, we have assumed that the isolated human clone is a fourth member of the VEGF family (hereinafter designated as VEGF-D).

[0008]

The present inventors analyzed the expression of the VEGF-D gene in various human tissues by northern blotting and found that the gene was expressed in lung, heart, small intestine, etc.

[0009]

The present invention relates to a novel gene belonging to the VEGF family. More specifically, it relates to

(1) A DNA described in SEQ ID NO: 2.

(2) A DNA hybridizing with the DNA described in SEQ ID NO: 2.

(3) A vector containing the DNA of (1) or (2).

(4) A transformant carrying the vector of (3).

(5) A protein encoded by the DNA of (1) or (2).

[0010]

[Mode for Carrying Out the Invention]

The VEGF-D of the present invention has high homology to VEGF-C and can be considered to be a fourth factor of the VEGF family. Since the major function of VEGF is vascular formation at the developmental stage and VEGF is considered to be involved in the pathological neovascularization associated with diabetes, rheumatoid arthritis, retinopathy, and the growth of solid tumors, VEGF-D is thought to have similar functions.

[0011]

VEGF-D of the present invention may be used for gene therapy by introducing the VEGF-D gene into the body of the patient with the VEGF-D deficiency, or expressing the gene in the body. An antisense DNA of the VEGF-D gene may also be used to inhibit the expression of the gene itself, thereby suppressing the pathological neovascularization.

[0012]

Among the many available methods to introduce the VEGF-D gene or its antisense DNA into the body, the retrovirus method, the liposome method, the cationic liposome method, and the adenovirus method are preferable.

[0013]

In order to express these genes in the body, the genes can be incorporated into a suitable vector and introduced into the body by

the retrovirus method, the liposome method, the cationic liposome method, or the adenovirus method. Although the vectors to be used are not particularly limited, such vectors as pAdexlcw and pZIPneo are preferable.

[0014]

The VEGF-D protein can also be used to heal wounds, promote collateral vessel formation, utilizing its angiogenic activity. Furthermore, an antibody against VEGF-D protein can be used as a therapeutic agent for pathological neovascularization.

[0015]

The VEGF-D protein used for the above purposes can be produced in a large quantity by inserting the VEGF-D gene into an appropriate vector and introducing the vector into a host. Vectors such as pGEMEX-1 (Promega) or pEF-BOS (Nucleic Acids Res. 1990, 18(17): p.5322) are preferable but not limited thereto. Suitable examples of the host into which the vector is introduced include *E. coli*, CHO cells, and COS cells.

[0016]

The VEGF-D protein expressed by the transformant in a large quantity can be purified by suitably combining purification treatments such as solubilization with a homogenizer or a sonicator, extraction by various buffers, solubilization or precipitation by acid or alkali, extraction or precipitation with organic solvents, salting out by ammonium sulfate and other agents, dialysis, ultrafiltration using membrane filters, gel filtration, ion exchange chromatography, reversed-phase chromatography, counter-current distribution chromatography, high-performance liquid chromatography, isoelectric focusing, gel electrophoresis, or affinity chromatography in which antibodies or receptors are immobilized.

[0017]

Antibodies against the VEGF-D protein can be prepared as polyclonal antibodies by immunizing rabbits, sheep, or other animals with the purified protein, or as monoclonal antibodies from the antibody-producing cells derived from immunized mice or rats. The antibodies thus obtained can be used as they are. When they are humanized, their immunogenicity can be effectively reduced. The methods of

humanizing the antibodies include the CDR graft method and the method of directly producing a human antibody. In the CDR Graft method, the antibody gene is cloned from the monoclonal antibody-producing cells and its antigenic determinant portion is transplanted into an existing human antibody. In the method of directly producing a human antibody, a mouse whose immune system has been replaced by the human immune system is immunized, similar to ordinary monoclonal antibodies.

[0018]

The VEGF-D protein or its antibody thus obtained can be administered into the body by subcutaneous injection or a similar method.

[0019]

A person skilled in the art could routinely isolate DNAs encoding analogous proteins based on the DNA sequence described in SEQ ID NO: 2 using hybridization techniques. The present invention includes a DNA that hybridizes with the VEGF-D DNA described in SEQ ID NO: 2 under specific conditions that it does not hybridize with VEGF-C DNA and a protein encoded by the DNA.

[0020]

[Examples]

[Example 1] Homology search by TFASTA method

The sequence CGPNKELDENTCQCVC was designed based on the consensus sequence found in the Balbiani ring 3 protein (BR3P) repeat at the C-terminus of VEGF-C. The entire ESTs and STS sequences in the Genbank database (as of 29 February 1996) were then searched by the TFASTA method (Person and Lipman, Proc. Natl. Acad. Sci. USA 85: 2444-2448 (1988)). The searching conditions used are shown below (Table 1).

[0021]

Table 1

Sequences	392,210
Symbols	135,585,305
Word Size	2
Gap creation penalty	12.0
Gap extension penalty	4.0

As a result, an EST (Accession No. H24828) that is considered to code the consensus sequence was found. The sequence is one of the ESTs registered by The WashU-Merck EST Project, and nine out of 16 amino acid residues were identical. Further searching for UniGene by NCBI based on this sequence revealed that five registered sequences (T64149, H24780, H24633, H24828, and T64277 (as of 1 March 1996)), including the above EST, were considered to be derived from the same gene. T64277 and T64149, as well as H24828 and H24780, are the combination of the 5' sequence and the 3' sequence of the same clones, and the length of the insert in both of these clones was 0.9 kb (Fig. 1).

[0022]

Translating the H24828 sequence into a protein sequence in a frame where homology is found suggested that this sequence codes 104 C-terminal amino acid residues. Comparing this amino acid sequence with the C-terminus of VEGF-C, 28 out of 104 amino acids (27%) were identical. Moreover, the amino acids that are important for maintaining the protein structure, such as cysteine and proline, were well conserved (Fig. 2).

[0023]

[Example 2] cDNA cloning from a library

Primers for 5' RACE and 3' RACE (5' RACE primer: 5'-AGGGATGGGGAACTTGGAACGCTGAAT-3', 3' RACE primer: 5'-GATCTAATCCAGCACCCCAAAAACTGC-3') were designed based on the EST (H24828) found in the search (Fig. 1). A double-stranded cDNA was synthesized from human lung-derived polyA$^+$ RNA using reverse transcriptase. PCR was then performed using Marathon-Ready cDNA, Lung

(Clontech), having an adapter cDNA ligated to both ends as a template cDNA, and using the above primer and adapter primer (AP-1 primer: 5'-CCATCCTAATACGACTCACTATAGGGC-3' (Fig. 1)) as primers. The above adapter cDNA contains the regions to which the adapter primers AP-1 and AP-2 hybridize. The PCR was performed in a manner such that the system was exposed to treatment at 94℃ for 1 min; five cycles of treatment at 94℃ for 30 sec and at 72℃ for 4 min; five cycles of treatment at 94℃ for 30 sec and at 70℃ for 4 min; then 25 cycles of treatment at 94℃ for 20 sec and at 68℃ for 4 min. (TaKaRa Ex Taq (Takara Shuzo) and the attached buffer were used as Taq polymerase instead of Advantage KlenTaq Polymerase Mix.) As a result, 1.5kb fragments were amplified at the 5' region and 0.9kb fragments at the 3' region. These fragments were cloned with the pCR-Direct Cloning System (Clontech), CR-TRAP Cloning System (GenHunter), and PT7Blue-T vector (Novagen). When the 5'-RACE fragment was cloned into the pCR-Direct vector, the fragment was amplified again using 5'-CTGGTTCGGCCCAGAACTTGGAACGCTGAATCA-3' and 5'-CTCGCTCGCCCACTAATACGACTCACTATAGG-3' as primers.

[0024]

[Example 3] Nucleotide sequence analysis

ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit with Amplitaq DNA Polymerase FS and 377 A DNA Sequencer (ABI) were used for DNA sequencing. The primers used are the primers in the vectors (5'-AATTAACCCTCACTAAAGGG-3', 5'-CCAGGGTTTTCCCAGTCACGAC-3'), AP-2 primer (5'-ACTCACTATAGGGCTCGAGCGGC-3'), and 10 primers in the sequence shown below (Table 2).

[0025]

Table 2

SQ1	5'-AAGTCTGGAGACCTGCT-3'
SQ2	5'-CAGCAGGTCTCCAGACT-3'
SQ3	5'-CGCACCCAAGGAATGGA-3'
SQ4	5'-TGACACCTGGCCATTCCA-3'
SQ5	5'-CATCAGATGGTAGTTCAT-3'
SQ6	5'-ATGCTGAGCGAGAGTCCATA-3'
SQ7	5'-CACTAGGTTTGCGGCAACTT-3'
SQ8	5'-GCTGTTGGCAAGCACTTACA-3'
SQ9	5'-GATCCATCCAGATCCCTGAA-3'
SQ10	5'-CAGATCAGGGCTGCTTCTA-3'

Determining the nucleotide sequence of the 1.5kb fragment at the 5'-side and the 0.9kb fragment at the 3'-side revealed that the sequence of the overlapping region was identical, confirming that 5'- and 3'-side cDNAs of the desired gene were obtained. Determining the entire nucleotide sequence of the cDNA revealed that this novel gene has the full length of 2 kb and can code a protein consisting of 354 amino acid residues (SEQ ID NO: 1 and SEQ ID NO: 2). Figure 1 shows the relation between this gene and the EST sequences registered in the Genbank database. Comparing the amino acid sequence with other VEGF family proteins revealed that the amino acids that are well conserved between family proteins are also conserved in this novel gene, and therefore this gene is obviously a new member of the VEGF family (Fig. 3). In Fig. 3, HSVEGF indicates human VEGF; HSVEGF-D, HSVEGF-C, and HSVEGF-B indicate human VEGF homologues (human VEGF-D, human VEGF-C, and human VEGF-B, respectively); HSPDGF-A indicates human PDGF-A; HSPDGF-B indicates human PDGF-B; and HSP1GF2 indicates human P1GF2. Since VEGF-D is highly homologous to VEGF-C that was cloned as the Flt4 ligand, it was presumed to be a ligand to a Flt-4-like receptor.

[0026]

Deducing the signal peptide cleavage site (Fig. 4b) by hydrophobicity plot (Fig. 4a) and the method of von Heijne (von Heijne,

G, Nucleic Acids Res. 14, 4683-4690(1986)), N-terminal 21 amino acid residues may be cleaved as signal peptides, and they may also undergo additional processing like VEGF-C.

[0027]

[Example 4] Northern blot analysis

A 1kb fragment, which had been cut out by digestion with EcoRI from the 5'-fragment subcloned into pCR-Direct vector, was labeled with [α-^{32}P]dCTP and used as a probe. Labeling was performed by random priming using Ready-to Go DNA labeling beads (Pharmacia). Hybridization was performed in ExpressHyb Hybridization Solution (Clontech) by the usual method using Multiple Tissue Northern (MTN) Blot-Human, Human II, Human Fetal, and Human Cell lines (Clontech). Significant expression was observed in lung, heart, and small intestine. Weak expression was observed in skeletal muscle, ovary, colon, and pancreas. The apparent molecular weight of the mRNA was 2.2 kb, and the cloned fragment seemed to be almost the full length of the gene.

[0028]

[Effect of the Invention]

In the present invention, a novel VEGF-D gene having high homology to the VEGF-C gene has been isolated. VEGF-D appears to be involved in the pathological neovascularization associated with diabetes, rheumatoid arthritis, and the growth of solid tumors as well as the normal neovascularization at the developmental stage. The gene of the present invention can be used to gene therapy for the VEGF-D deficiency, and the VEGF-D protein, which is obtained by expressing the gene of the present invention, can be used to heal wounds and promote collateral vessel formation. The inhibitors against the VEGF-D protein can be used as a novel anticancer agent.

[0029]

[Sequence listing]

SEQ ID NO: 1:

SEQUENCE LENGTH: 2004

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

TISSUE TYPE: lung

FEATURE:

NAME/KEY: CDS

LOCATION: 403 .. 1464

IDENTIFICATION METHOD: E

SEQUENCE DESCRIPTION: SEQ ID NO: 1

```
CCAGCTTTCT GTARCTGTAA GCATTGGTGG CCACACCACC TCCTTACAAA GCAACTAGAA      60
CCTGCGGCAT ACATTGGAGA GATTTTTTTA ATTTTCTGGA CAYGAAGTAA ATTTAGAGTG     120
CTTTCYAATT TCAGGTAGAA GACATGTCCA CCTTCTGATT ATTTTTGGAG AACATTTTGA     180
TTTTTTTCAT CTCTCTCTCC CCACCCCTAA GATTGTGCAA AAAAAGCGTA CCTTGCCTAA     240
TTGAAATAAT TTCATTGGAT TTTGATCAGA ACTGATCATT TGGTTTTCTG TGTGAAGTTT     300
TGAGGTTTCA AACTTTCCTT CTGGAGAATG CCTTTTGAAA CAATTTTCTC TAGCTGCCTG     360
ATGTCAACTG CTTAGTAATC AGTGGATATT GAAATATTCA AA ATG TAC AGA GAG         414
                                               Met Tyr Arg Glu
                                               1
TGG GTA GTG GTG AAT GTT TTC ATG ATG TTG TAC GTC CAG CTG GTG CAG       462
Trp Val Val Val Asn Val Phe Met Met Leu Tyr Val Gln Leu Val Gln
5                   10                  15                  20
GGC TCC AGT AAT GAA CAT GGA CCA GTG AAG CGA TCA TCT CAG TCC ACA       510
Gly Ser Ser Asn Glu His Gly Pro Val Lys Arg Ser Ser Gln Ser Thr
                25                  30                  35
TTG GAA CGA TCT GAA CAG CAG ATC AGG GCT GCT TCT AGT TTG GAG GAA       558
Leu Glu Arg Ser Glu Gln Gln Ile Arg Ala Ala Ser Ser Leu Glu Glu
            40                  45                  50
CTA CTT CGA ATT ACT CAC TCT GAG GAC TGG AAG CTG TGG AGA TGC AGG       606
Leu Leu Arg Ile Thr His Ser Glu Asp Trp Lys Leu Trp Arg Cys Arg
```

```
        55                  60                  65
CTG AGG CTC AAA AGT TTT ACC AGT ATG GAC TCT CGC TCA GCA TCC CAT      654
Leu Arg Leu Lys Ser Phe Thr Ser Met Asp Ser Arg Ser Ala Ser His
    70                  75                  80
CGG TCC ACT AGG TTT GCG GCA ACT TTC TAT GAC ATT GAA ACA CTA AAA      702
Arg Ser Thr Arg Phe Ala Ala Thr Phe Tyr Asp Ile Glu Thr Leu Lys
85                  90                  95                  100
GTT ATA GAT GAA GAA TGG CAA AGA ACT CAG TGC AGC CCT AGA GAA ACG      750
Val Ile Asp Glu Glu Trp Gln Arg Thr Gln Cys Ser Pro Arg Glu Thr
                105                 110                 115
TGC GTG GAG GTG GCC AGT GAG CTG GGG AAG AGT ACC AAC ACA TTC TTC      798
Cys Val Glu Val Ala Ser Glu Leu Gly Lys Ser Thr Asn Thr Phe Phe
            120                 125                 130
AAG CCC CCT TGT GTG AAC GTG TTC CGA TGT GGT GGC TGT TGC AAT GAA      846
Lys Pro Pro Cys Val Asn Val Phe Arg Cys Gly Gly Cys Cys Asn Glu
        135                 140                 145
GAG AGC CTT ATC TGT ATG AAC ACC AGC ACC TCG TAC ATT TCC AAA CAG      894
Glu Ser Leu Ile Cys Met Asn Thr Ser Thr Ser Tyr Ile Ser Lys Gln
    150                 155                 160
CTC TTT GAG ATA TCA GTG CCT TTG ACA TCA GTA CCT GAA TTA GTG CCT      942
Leu Phe Glu Ile Ser Val Pro Leu Thr Ser Val Pro Glu Leu Val Pro
165                 170                 175                 180
GTT AAA GTT GCC AAT CAT ACA GGT TGT AAG TGC TTG CCA ACA GCC CCC      990
Val Lys Val Ala Asn His Thr Gly Cys Lys Cys Leu Pro Thr Ala Pro
                185                 190                 195
CGC CAT CCA TAC TCA ATT ATC AGA AGA TCC ATC CAG ATC CCT GAA GAA     1038
Arg His Pro Tyr Ser Ile Ile Arg Arg Ser Ile Gln Ile Pro Glu Glu
            200                 205                 210
GAT CGC TGT TCC CAT TCC AAG AAA CTC TGT CCT ATT GAC ATG CTA TGG     1086
Asp Arg Cys Ser His Ser Lys Lys Leu Cys Pro Ile Asp Met Leu Trp
        215                 220                 225
GAT AGC AAC AAA TGT AAA TGT GTT TTG CAG GAG GAA AAT CCA CTT GCT     1134
Asp Ser Asn Lys Cys Lys Cys Val Leu Gln Glu Glu Asn Pro Leu Ala
    230                 235                 240
GGA ACA GAA GAC CAC TCT CAT CTC CAG GAA CCA GCT CTC TGT GGG CCA     1182
Gly Thr Glu Asp His Ser His Leu Gln Glu Pro Ala Leu Cys Gly Pro
```

```
245                 250                 255                 260
CAC ATG ATG TTT GAC GAA GAT CGT TGC GAG TGT GTC TGT AAA ACA CCA     1230
His Met Met Phe Asp Glu Asp Arg Cys Glu Cys Val Cys Lys Thr Pro
                265                 270                 275
TGT CCC AAA GAT CTA ATC CAG CAC CCC AAA AAC TGC AGT TGC TTT GAG     1278
Cys Pro Lys Asp Leu Ile Gln His Pro Lys Asn Cys Ser Cys Phe Glu
            280                 285                 290
TGC AAA GAA AGT CTG GAG ACC TGC TGC CAG AAG CAC AAG CTA TTT CAC     1326
Cys Lys Glu Ser Leu Glu Thr Cys Cys Gln Lys His Lys Leu Phe His
        295                 300                 305
CCA GAC ACC TGC AGC TGT GAG GAC AGA TGC CCC TTT CAT ACC AGA CCA     1374
Pro Asp Thr Cys Ser Cys Glu Asp Arg Cys Pro Phe His Thr Arg Pro
    310                 315                 320
TGT GCA AGT GGC AAA ACA GCA TGT GCA AAG CAT TGC CGC TTT CCA AAG     1422
Cys Ala Ser Gly Lys Thr Ala Cys Ala Lys His Cys Arg Phe Pro Lys
325                 330                 335                 340
GAG AAA AGG GCT GCC CAG GGG CCC CAC AGC CGA AAG AAT CCT             1464
Glu Lys Arg Ala Ala Gln Gly Pro His Ser Arg Lys Asn Pro
                345                 350
TGATTCAGCG TTCCAAGTTC CCCATCCCTG TCATTTTTAA CAGCATGCTG CTTTGCCAAG   1524
TTGCTGTCAC TGTTTTTTTC CCAGGTGTTA AAAAAAAAAT CCATTTTACA CAGCACCACA   1584
GTGAATCCAG ACCAACCTTC CATTCACACC AGCTAAGGAG TCCCTGGTTC ATTGATGGAT   1644
GTCTTCTAGC TGCAGATGCC TCTGCGCACC AAGGAATGGA GAGGAGGGGA CCCATGTAAT   1704
CCTTTTGTTT AGTTTTGTTT TTGTTTTTTG GTGAATGAGA AAGGTGTGCT GGTCATGGAA   1764
TGGCAGGTGT CATATGACTG ATTACTCAGA GCAGATGAGG AAAACTGTAG TCTCTGAGTC   1824
CTTTGCTAAT CGCAACTCTT GTGAATTATT CTGATTCTTT TTTATGCAGA ATTTGATTCG   1884
TATGATCAGT ACTGACTTTC TGATTACTGT CCAGCTTATA GTCTTCCAGT TTAATGAACT   1944
ACCATCTGAT GTTTCATATT TAAGTGTATT TAAAGAAAAT AAACACCATT ATTCAAGTCT   2004
```

SEQ ID NO: 2:

SEQENCE LENGTH: 1062

SEQENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens
TISSUE TYPE: lung
FEATURE:
NAME/KEY: CDS
LOCATION: 1 .. 1062
IDENTIFICATION METHOD: E
SEQUENCE DESCRIPTION: SEQ ID NO: 2

```
ATG TAC AGA GAG TGG GTA GTG GTG AAT GTT TTC ATG ATG TTG TAC GTC      48
Met Tyr Arg Glu Trp Val Val Val Asn Val Phe Met Met Leu Tyr Val
1               5                  10                  15
CAG CTG GTG CAG GGC TCC AGT AAT GAA CAT GGA CCA GTG AAG CGA TCA      96
Gln Leu Val Gln Gly Ser Ser Asn Glu His Gly Pro Val Lys Arg Ser
            20                  25                  30
TCT CAG TCC ACA TTG GAA CGA TCT GAA CAG CAG ATC AGG GCT GCT TCT     144
Ser Gln Ser Thr Leu Glu Arg Ser Glu Gln Gln Ile Arg Ala Ala Ser
        35                  40                  45
AGT TTG GAG GAA CTA CTT CGA ATT ACT CAC TCT GAG GAC TGG AAG CTG     192
Ser Leu Glu Glu Leu Leu Arg Ile Thr His Ser Glu Asp Trp Lys Leu
    50                  55                  60
TGG AGA TGC AGG CTG AGG CTC AAA AGT TTT ACC AGT ATG GAC TCT CGC     240
Trp Arg Cys Arg Leu Arg Leu Lys Ser Phe Thr Ser Met Asp Ser Arg
65                  70                  75                  80
TCA GCA TCC CAT CGG TCC ACT AGG TTT GCG GCA ACT TTC TAT GAC ATT     288
Ser Ala Ser His Arg Ser Thr Arg Phe Ala Ala Thr Phe Tyr Asp Ile
                85                  90                  95
GAA ACA CTA AAA GTT ATA GAT GAA GAA TGG CAA AGA ACT CAG TGC AGC     336
Glu Thr Leu Lys Val Ile Asp Glu Glu Trp Gln Arg Thr Gln Cys Ser
            100                 105                 110
CCT AGA GAA ACG TGC GTG GAG GTG GCC AGT GAG CTG GGG AAG AGT ACC     384
Pro Arg Glu Thr Cys Val Glu Val Ala Ser Glu Leu Gly Lys Ser Thr
        115                 120                 125
AAC ACA TTC TTC AAG CCC CCT TGT GTG AAC GTG TTC CGA TGT GGT GGC     432
Asn Thr Phe Phe Lys Pro Pro Cys Val Asn Val Phe Arg Cys Gly Gly
    130                 135                 140
TGT TGC AAT GAA GAG AGC CTT ATC TGT ATG AAC ACC AGC ACC TCG TAC     480
Cys Cys Asn Glu Glu Ser Leu Ile Cys Met Asn Thr Ser Thr Ser Tyr
```

```
145                 150                 155                 160
ATT TCC AAA CAG CTC TTT GAG ATA TCA GTG CCT TTG ACA TCA GTA CCT      528
Ile Ser Lys Gln Leu Phe Glu Ile Ser Val Pro Leu Thr Ser Val Pro
                165                 170                 175
GAA TTA GTG CCT GTT AAA GTT GCC AAT CAT ACA GGT TGT AAG TGC TTG      576
Glu Leu Val Pro Val Lys Val Ala Asn His Thr Gly Cys Lys Cys Leu
            180                 185                 190
CCA ACA GCC CCC CGC CAT CCA TAC TCA ATT ATC AGA AGA TCC ATC CAG      624
Pro Thr Ala Pro Arg His Pro Tyr Ser Ile Ile Arg Arg Ser Ile Gln
        195                 200                 205
ATC CCT GAA GAA GAT CGC TGT TCC CAT TCC AAG AAA CTC TGT CCT ATT      672
Ile Pro Glu Glu Asp Arg Cys Ser His Ser Lys Lys Leu Cys Pro Ile
    210                 215                 220
GAC ATG CTA TGG GAT AGC AAC AAA TGT AAA TGT GTT TTG CAG GAG GAA      720
Asp Met Leu Trp Asp Ser Asn Lys Cys Lys Cys Val Leu Gln Glu Glu
225                 230                 235                 240
AAT CCA CTT GCT GGA ACA GAA GAC CAC TCT CAT CTC CAG GAA CCA GCT      768
Asn Pro Leu Ala Gly Thr Glu Asp His Ser His Leu Gln Glu Pro Ala
                245                 250                 255
CTC TGT GGG CCA CAC ATG ATG TTT GAC GAA GAT CGT TGC GAG TGT GTC      816
Leu Cys Gly Pro His Met Met Phe Asp Glu Asp Arg Cys Glu Cys Val
            260                 265                 270
TGT AAA ACA CCA TGT CCC AAA GAT CTA ATC CAG CAC CCC AAA AAC TGC      864
Cys Lys Thr Pro Cys Pro Lys Asp Leu Ile Gln His Pro Lys Asn Cys
        275                 280                 285
AGT TGC TTT GAG TGC AAA GAA AGT CTG GAG ACC TGC TGC CAG AAG CAC      912
Ser Cys Phe Glu Cys Lys Glu Ser Leu Glu Thr Cys Cys Gln Lys His
    290                 295                 300
AAG CTA TTT CAC CCA GAC ACC TGC AGC TGT GAG GAC AGA TGC CCC TTT      960
Lys Leu Phe His Pro Asp Thr Cys Ser Cys Glu Asp Arg Cys Pro Phe
305                 310                 315                 320
CAT ACC AGA CCA TGT GCA AGT GGC AAA ACA GCA TGT GCA AAG CAT TGC     1008
His Thr Arg Pro Cys Ala Ser Gly Lys Thr Ala Cys Ala Lys His Cys
                325                 330                 335
CGC TTT CCA AAG GAG AAA AGG GCT GCC CAG GGG CCC CAC AGC CGA AAG     1056
Arg Phe Pro Lys Glu Lys Arg Ala Ala Gln Gly Pro His Ser Arg Lys
```

340 345 350

AAT CCT 1062

Asn Pro

[Brief Description of the Drawings]

[Fig. 1]

Figure 1 shows the relationship among the VEGF-D gene, the EST sequences, and the primers used for cloning.

[Fig. 2]

Figure 2 compares the amino acid sequences of EST (H24828) and VEGF-C.

[Fig. 3]

Figure 3 compares the amino acid sequences deduced from the VEGF-D gene and from the known genes of the VEGF family proteins.

[Fig. 4]

Figure 4a shows the hydrophobicity plot of VEGF-D. Figure 4b shows the prediction of the cleavage site of the VEGF-D signal peptide.

[Document Name] Abstract

[Abstract]

[Objective] An objective of the present invention is to isolate a novel gene of the VEGF family.

[Means of Solution] PCR was performed to amplify cDNA using primers designed based on the sequence of EST that is assumed to have homology to the C-terminal region of the VEGF-C gene, a VEGF family gene. The amplified cDNA was sequenced, and the amino acid sequence deduced therefrom revealed that the amino acid sequence had high homology to that of VEGF-C as a whole. The expression of the VEGF-D gene in various human tissues was examined by northern blotting, and the gene was found to be expressed in lung, heart, and small intestine.

[Selected Drawing] Figure 3